

## ABSTRACT OF THE DISCLOSURE

A novel seed coat specific peroxidase genomic sequence is characterized and presented. Adjacent DNA regulatory regions have also been characterized. The seed coat peroxidase is translated as a 352 amino acid precursor protein of 38 kDa comprising a 26 amino acid signal sequence which when cleaved results in a 35 kDa protein. Plants containing a dominant *Ep* allele accumulate large amounts of peroxidase in the hourglass cells of the subepidermis. Homozygous recessive *epep* genotypes do not accumulate peroxidase in the hourglass cells and are much reduced in total seed coat peroxidase activity. Probes derived from the cDNA, or genomic DNA can be used to detect polymorphisms that distinguished *EpEp* and *epep* genotypes. Cosegregation of the polymorphisms in an  $F_2$  population from a cross of *EpEp* and *epep* plants shows that the *Ep* locus encodes the seed coat peroxidase protein. Comparison of *Ep* and *ep* alleles indicates that the recessive gene lacks 87 bp of sequence encompassing the translation start codon. The heterologous expression, as well as vectors and hosts to be used for the expression of the seed coat peroxidase, are also disclosed. The seed-specific DNA regulatory region may be used to control expression of genes of interest such as i) genes encoding herbicide resistance, or ii) biological control of insects or pathogens (e.g. *B. thuringiensis*), or iii) viral coat proteins to protect against viral infections, or iv) proteins of commercial interest (e.g. pharmaceutical), and v) proteins that alter the nutritive value, taste, or processing of seeds.